Hash Function Standards and Practices in Global Digital Forensics Laboratories

I. Introduction: The Foundational Role of Hash Functions in Digital Forensics

A. Defining Hash Functions and Their Significance

In the realm of digital forensics, establishing the integrity and authenticity of electronic evidence is paramount. Central to this objective are cryptographic hash functions. A hash function is essentially a mathematical algorithm that takes an input (a message or file of any size) and produces a fixed-size string of characters, known as a hash value, message digest, or digital fingerprint. This process maps data of arbitrary length to a condensed representation of a defined length.

The defining characteristic of cryptographic hash functions is their one-way nature; it is computationally infeasible to derive the original input data solely from its hash value. This property is crucial for security applications. Furthermore, robust hash functions exhibit several key cryptographic properties essential for forensic reliability :

1. **Preimage Resistance:** Given a specific hash value, it is computationally infeasible to find an input message that produces that hash value. This prevents an attacker from crafting a fraudulent message to match a known hash.
2. **Second Preimage Resistance:** Given a specific input message, it is computationally infeasible to find a *different* input message that produces the exact same hash value. This property is vital for ensuring that an original piece of evidence cannot be replaced with altered content that shares the same hash.
3. **Collision Resistance:** It is computationally infeasible to find *any* two distinct input messages that produce the same hash value. While related to second preimage resistance, this property is stricter, as it doesn't require starting with a specific message.

The subtle but important distinctions between these properties, particularly between collision resistance and second preimage resistance, have significant implications in the forensic context. While cryptographic standards bodies often focus heavily on collision resistance due to its importance in applications like digital signatures, the forensic community has sometimes placed greater emphasis on second preimage resistance when evaluating algorithms for evidence integrity verification. This difference arises because the typical forensic scenario involves verifying the integrity of a *known* piece of evidence against its original hash, making the ability to find a second preimage for that specific evidence the more direct threat, rather than simply finding any two arbitrary files that collide.

B. Primary Forensic Applications

Hash functions serve several critical roles within the digital forensic workflow:

1. **Integrity Verification:** This is arguably the most fundamental application. By calculating the hash value of digital evidence (such as a file, a disk partition, or a complete forensic

image) at the time of acquisition and comparing it to hash values calculated at subsequent stages (e.g., after creating a working copy, during analysis, before presentation in court), examiners can verify that the evidence has not been altered. Any change to the input data, even a single bit, will result in a drastically different hash value. This process provides a verifiable "digital fingerprint" that forms a cornerstone of the chain of custody for digital evidence, ensuring that the evidence presented is the same as what was originally collected. The standard procedure involves computing a hash of the original evidence, creating a working copy, computing a hash of the working copy, and confirming the two hashes match.

2. **File Identification and Classification:** Hash values are widely used to quickly identify known files. Forensic laboratories often compare the hash values of files found on suspect media against large databases of known hashes. The most prominent example is the National Software Reference Library (NSRL) maintained by the U.S. National Institute of Standards and Technology (NIST), which contains hash values of known operating system files, common applications, and other legitimate software. Comparing hashes allows examiners to quickly filter out irrelevant system files (known-good) or identify known-bad files, such as malware, illicit images, or contraband data, without needing to manually examine every file. This technique significantly aids in data reduction, triage, and focusing investigative efforts.
3. **Digital Signatures (Contextual Relevance):** While not solely a forensic application, hash functions are an indispensable component of digital signatures. A digital signature typically involves hashing the data to be signed and then encrypting the hash value with the signatory's private key. The recipient can verify the signature using the signatory's public key, confirming both the data's integrity (it hasn't changed since signing) and the signatory's identity (non-repudiation). In a forensic context, verifying digital signatures on documents or communications can be crucial. However, the security requirements for hash functions used in digital signatures are particularly stringent, especially concerning collision resistance, as an attacker finding a collision could potentially substitute a fraudulent document for a legitimate one under the same signature.

C. The Need for Standardization and Best Practices

The admissibility of digital evidence in legal proceedings hinges on the demonstration of sound, validated, and reliable forensic processes. The use of scientifically accepted methodologies ensures that the evidence presented is trustworthy and that the conclusions drawn are scientifically supported. International and national bodies play a crucial role in establishing these benchmarks. Organizations such as NIST in the United States, the Scientific Working Group on Digital Evidence (SWGDE), the European Network of Forensic Science Institutes (ENFSI), and INTERPOL develop standards, best practice manuals, and guidelines that shape the procedures used in digital forensic laboratories worldwide.

However, the rapidly evolving nature of technology presents a continuous challenge. Cryptographic hash functions, once considered secure, can become vulnerable to new analytical techniques and increased computing power. This necessitates periodic updates to standards and practices. A significant challenge lies in balancing the adoption of newer, more secure cryptographic algorithms (like the SHA-2 and SHA-3 families) with the practical constraints faced by forensic laboratories. These constraints include the need to process legacy data hashed with older algorithms, the time and resources required to validate new tools and software implementing updated algorithms, and ensuring interoperability between different tools

and organizations.
The landscape of hash function standards reflects this dynamic. Foundational cryptographic standards, often driven by bodies like NIST with a focus on broad information security for government and industry, tend to push aggressively for migration away from algorithms with known weaknesses. Concurrently, applied forensic practice guidelines from groups like SWGDE, NFI, ENFSI, and INTERPOL aim to provide practical recommendations for examiners, sometimes acknowledging operational realities and offering nuanced positions on the use of older algorithms for specific forensic tasks while still promoting the adoption of stronger methods. Understanding this interplay between rigorous cryptographic standards and applied forensic practice is essential for interpreting the current state of hash function usage in the field.

II. Dominant Hash Standards and Recommendations: The Move Towards SHA-2 and SHA-3

The selection and use of hash functions in digital forensics are heavily influenced by standards and recommendations issued by authoritative bodies. While practices may vary slightly by jurisdiction and laboratory, the guidelines set forth by organizations like NIST, ENFSI, INTERPOL, and national institutes like NFI provide a framework widely adopted globally. The clear trend is a move away from legacy algorithms towards the more secure SHA-2 and SHA-3 families.

A. The NIST Framework (USA Focus, Global Influence)

The U.S. National Institute of Standards and Technology (NIST) plays a pivotal role in establishing cryptographic standards, particularly through its Federal Information Processing Standards (FIPS) and Special Publications (SP). While primarily mandated for U.S. Federal agencies , NIST standards are highly influential and often adopted by industry and governments worldwide.

- **FIPS 180-4 (Secure Hash Standard):** This standard specifies the approved hash algorithms for federal use. It currently includes the legacy SHA-1 algorithm and the widely adopted SHA-2 family, which comprises SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224, and SHA-512/256.
- **FIPS 202 (SHA-3 Standard):** Resulting from a public competition seeking a next-generation hash algorithm, FIPS 202 specifies the SHA-3 family. Based on the Keccak algorithm, this family includes four fixed-length hash functions (SHA3-224, SHA3-256, SHA3-384, SHA3-512) and two Extendable-Output Functions (XOFs), SHAKE128 and SHAKE256.
- **NIST Policy Evolution (SHA-1 Deprecation):** NIST has pursued a deliberate and increasingly stringent policy to phase out SHA-1 due to mounting cryptographic weaknesses.
    - As early as 2006, NIST recommended using SHA-2 for new applications and mandated that federal agencies stop using SHA-1 for digital signatures and collision-resistance applications after 2010.
    - In 2011, NIST formally deprecated the use of SHA-1, and disallowed its use for digital signatures entirely at the end of 2013, citing both theoretical attacks (like the Wang et al. attack) and the increasing feasibility of brute-force attacks.

    - Following the practical demonstration of a SHA-1 collision in 2017 , NIST intensified its recommendations, urging federal agencies to migrate away from SHA-1 for *all* applications as soon as possible.
    - The final deadline has been set: After **December 31, 2030**, SHA-1 will be completely removed from the FIPS 180 standard, and cryptographic modules validated under FIPS 140 that include SHA-1 will be moved to a historical list, rendering them unsuitable for federal procurement. This deadline exerts significant pressure on vendors of cryptographic modules, including those embedded in forensic tools, to update their products.
- **NIST Recommendations for SHA-2/SHA-3:**
    - The SHA-2 family of hash functions remains fully approved for all applications requiring secure hash algorithms within the federal government.
    - NIST specifically encourages the implementation of **SHA-256 as a minimum** standard for any application requiring interoperability, reflecting its widespread adoption and robust security profile.
    - The SHA-3 family (fixed-length variants) is also approved for all applications as a secure alternative to SHA-2. NIST has clarified that, at present, there is **no requirement or plan to mandate a transition *from* SHA-2 *to* SHA-3**. SHA-2 is still considered secure for the foreseeable future.
- **Security Strengths:** FIPS 180-4 and related NIST documents define the expected security strengths (in bits) for approved hash functions against different types of attacks. For example, SHA-256 provides 128 bits of collision resistance and 256 bits of preimage resistance, while SHA-512 provides 256 bits of collision resistance and 512 bits of preimage resistance. In contrast, SHA-1's collision resistance is considered less than 80 bits, falling below acceptable security margins for many applications. These defined strengths guide the selection of appropriate algorithms based on the security needs of the application.

NIST's clear roadmap for SHA-1 deprecation and its endorsement of SHA-2 and SHA-3 set a strong precedent. While directly binding only on U.S. federal agencies, this framework significantly influences international practices, software development, and the capabilities embedded within commercial forensic tools, pushing the entire digital forensics field towards stronger cryptographic foundations.

B. International Guidelines and Practices

Beyond NIST, several international bodies and national forensic institutes provide guidance that reflects both global best practices and regional considerations.

- **ENFSI (European Network of Forensic Science Institutes):** ENFSI focuses on promoting harmonization and establishing Best Practice Manuals (BPMs) across various forensic disciplines, including digital technology. The ENFSI BPM for the Forensic Examination of Digital Technology (issued in 2015) underscores the critical importance of maintaining the integrity of acquired digital evidence throughout the forensic process. It mandates establishing an auditable link between the acquired data and the original source and emphasizes the need for rigorous function verification of forensic tools, explicitly mentioning hash calculation as a function whose results must be trustworthy and potentially subject to formal verification. While the 2015 BPM does not mandate the use of specific hash algorithms like SHA-256 in the available excerpts, its strong focus on integrity, auditability, tool verification, and risk mitigation clearly implies the necessity of

employing robust, cryptographically sound hashing algorithms that align with contemporary security standards. The choice might be left to member labs based on current cryptographic advice and validation status.

- **INTERPOL:** As a global law enforcement organization, INTERPOL provides guidelines aimed at fostering international cooperation and capacity building. The INTERPOL Global Guidelines for Digital Forensics Laboratories (2019) explicitly recommend using hashing algorithms to verify the integrity of forensic images, ensuring the copy matches the original exhibit. The guidelines specifically cite **SHA-256** as an example of a hashing algorithm used in digital forensics, while also acknowledging that various algorithms are employed in the field and that most forensic tools incorporate hashing capabilities. INTERPOL's guidelines for first responders also highlight the importance of evidence preservation and integrity , often referencing other established guides like those from the Council of Europe.
- **NFI (Netherlands Forensic Institute):** The NFI represents a proactive national institute. Their 2018 technical supplement on hash values indicates that their Digital and Biometric Traces (DBS) division made a strategic decision to **switch to SHA-256 as early as mid-2010**. This move was a preventive measure based on emerging research highlighting potential weaknesses in the second preimage resistance of MD5 and SHA-1, even before practical attacks were widely demonstrated. While proactively adopting SHA-256 for their own work, the NFI acknowledged the practical necessity of verifying MD5 and SHA-1 hashes received from chain partners using older software. As of 2018, they considered SHA-256 sufficiently secure for their purposes and had not yet transitioned to SHA-3, deeming its adoption not yet common in the forensic community. The NFI's development and use of the Hansken Digital Forensics as a Service (DFaaS) platform further suggests a commitment to standardized, robust processing pipelines, likely incorporating SHA-256 or stronger hashing.
- **BKA (Bundeskriminalamt, Germany):** Germany's Federal Criminal Police Office operates a significant forensic science institute and engages in complex cybercrime investigations. Documentation indicates that their computer forensics activities are conducted under accredited quality management systems (DIN EN ISO/IEC 17025 and 17020). Such accreditation necessitates the use of validated and scientifically sound methods. While specific internal BKA standards for hash algorithms are not detailed in the provided materials, their commitment to quality management, role as a central agency interacting with state and international partners , and involvement in technical discussions (e.g., related to hashing in travel documents) strongly suggest alignment with current international best practices. This would almost certainly involve the use of SHA-2 family algorithms (like SHA-256) for generating new hashes in current investigations, coupled with the capability to analyze evidence containing legacy hashes.
- **FBI (USA):** As a major U.S. federal law enforcement agency, the FBI's practices are directly influenced by NIST standards and guidelines. Internal procedures and systems development would adhere to FIPS requirements, mandating the use of SHA-2 or SHA-3 for applications requiring cryptographic hashing. Joint guidance issued by the FBI with other agencies explicitly recommends using SHA-256 for verifying the integrity of network device configurations. The FBI Laboratory emphasizes the importance of evidence integrity and uses standard forensic processes that include hash verification. While FBI labs must be equipped to analyze evidence containing any type of hash encountered, their own generated hashes and internal systems would align with the NIST-driven migration towards SHA-2 and SHA-3.

- **Europol:** Europol's role involves facilitating cooperation and information exchange among EU member states in combating serious international crime, including cybercrime. They operate platforms for data sharing and analysis, including a decryption platform to assist national authorities. Given its mandate to handle sensitive data from diverse sources and ensure the integrity of its processes, Europol's internal practices would necessitate robust security measures, including the use of strong, standardized hash functions (likely SHA-2 or SHA-3) for data integrity and internal systems. However, like other central bodies, Europol must also be capable of ingesting and processing evidence submitted by member states, which may utilize varying hashing standards depending on national practices and the age of the evidence. The provided documents emphasize legal frameworks, encryption challenges, and AI applications more than specific internal hashing standards.

Overall, the international landscape shows a strong consensus on the need for robust hashing for evidence integrity, with SHA-256 emerging as a de facto minimum standard or common recommendation, driven by NIST's influence and the known weaknesses of older algorithms. However, practical implementation varies, and the need to handle legacy data persists.

III. The Status of Legacy Algorithms (MD5 & SHA-1) in Digital Forensics

Despite the clear trend towards SHA-2 and SHA-3, the legacy algorithms MD5 and SHA-1 continue to be encountered and, in some limited contexts, used within the digital forensics community. Understanding their status requires acknowledging their known cryptographic weaknesses while appreciating the specific ways they have been employed forensically.

A. Acknowledged Cryptographic Weaknesses

Both MD5 and SHA-1 are considered cryptographically broken for applications that demand strong collision resistance.

- **MD5:** Developed in the early 1990s , MD5 produces a 128-bit hash value. Significant weaknesses were identified, and the first practical collision attack was demonstrated by Wang et al. in 2004. This means it is feasible to intentionally create two different files that produce the same MD5 hash. Consequently, MD5 is widely regarded as insecure for applications like digital signatures, software integrity verification (where malicious modification is a threat), and password hashing (due to its speed facilitating brute-force attacks). Major software vendors and organizations have deprecated its use for security-critical functions. Despite its weaknesses, its speed remains an advantage in some non-security-critical contexts.
- **SHA-1:** Introduced as a FIPS standard in 1995 , SHA-1 generates a 160-bit hash value. Theoretical attacks suggesting weaknesses emerged as early as 2005, again pioneered by Wang et al.. These theoretical attacks prompted NIST to begin phasing out SHA-1. The concerns were realized in February 2017 when researchers from CWI Amsterdam and Google announced the first practical SHA-1 collision, known as the SHAttered attack. They successfully created two different PDF files with the same SHA-1 hash value. Further research has refined these attacks, including demonstrating chosen-prefix collisions in 2019, which offer attackers more flexibility.
- **Implications:** The existence of practical collision attacks fundamentally undermines the trustworthiness of MD5 and SHA-1 for any application where collision resistance is

essential. The primary risk is that an adversary could create a malicious file (e.g., malware, a fraudulent contract) that shares the same hash value as a legitimate file. This could allow the malicious file to pass an integrity check based solely on the compromised hash algorithm, or to be accepted under a digital signature generated for the legitimate file.

B. The Forensic Distinction: Collision vs. (Second) Preimage Resistance

While the collision vulnerabilities of MD5 and SHA-1 are undisputed, some segments of the forensic community have historically maintained that these algorithms could still be acceptable for the specific task of *evidence integrity verification*. This position stems from the crucial difference between collision resistance and second preimage resistance.
As defined earlier, collision resistance concerns finding *any* two distinct inputs M_1 and M_2 such that Hash(M_1) = Hash(M_2). Second preimage resistance, on the other hand, concerns being given a specific input M_1 and its hash H_1 = Hash(M_1), and then finding a *different* input M_2 such that Hash(M_2) is also equal to H_1.
In the context of verifying the integrity of a specific piece of digital evidence (e.g., a forensic image file E), the examiner calculates Hash(E) at the time of acquisition. Later, they recalculate the hash of the working copy or the evidence itself to ensure no modifications have occurred. The direct threat model here is an adversary attempting to alter the original evidence E to create a modified version E' such that Hash(E') = Hash(E). This is precisely the second preimage problem.
For a considerable period after collision attacks on MD5 and SHA-1 were known, practical attacks demonstrating the ability to find second preimages for arbitrary files were not available or were considered computationally infeasible for typical forensic scenarios. Finding a collision generally requires significant computational effort and often results in files containing specific "collision-causing" artifacts, which might be detectable during a thorough forensic examination. Finding a second preimage for a specific, pre-existing file without such artifacts was seen as a much harder problem. It was this distinction – the perceived resilience against second preimage attacks despite known collision vulnerabilities – that formed the technical basis for some forensic bodies continuing to deem MD5 and SHA-1 acceptable *solely* for verifying the integrity of a specific piece of evidence against its previously calculated hash.

C. Guidance from Forensic Bodies (SWGDE, NFI): Continued Use with Caveats

Reflecting the distinction discussed above, guidance from practitioner-focused forensic bodies has sometimes diverged from the stricter cryptographic recommendations.

- **SWGDE Position (2019):** The Scientific Working Group on Digital Evidence issued a position statement in September 2019 explicitly addressing the use of MD5 and SHA-1. It stated: "While SWGDE promotes the adoption of SHA2 and SHA3 by vendors and practitioners, the MD5 and SHA1 algorithms remain acceptable for integrity verification and file identification applications in digital forensics. Because of known limitations of the MD5 and SHA1 algorithms, only SHA2 and SHA3 are appropriate for digital signatures and other security applications.". This position clearly differentiates between the integrity verification task (where second preimage resistance is key) and security applications like

signatures (where collision resistance is paramount). While technically defensible based on the state of attacks *at that time* regarding second preimages, this position was criticized by some cryptographers as "bad form" for endorsing algorithms with known significant weaknesses, potentially hindering migration efforts.

- **NFI Supplement (2018):** The Netherlands Forensic Institute's 2018 supplement similarly argued that MD5 and SHA-1 were "still perfectly usable for integrity control... and identification and classification" provided their one-way (preimage) and target collision (second preimage) resistance properties held. They explicitly stated that the cracking of *random* collision resistance was less critical for these specific forensic tasks. However, demonstrating a more cautious approach aligned with cryptographic advancements, NFI also noted their preventive switch to SHA-256 back in 2010 due to the *potential* for future second preimage attacks becoming feasible.
- **Practical Realities:** The persistence of MD5 and SHA-1 in forensic workflows is often linked to practical factors. Legacy systems and forensic tools may only support these older algorithms, and updating software requires significant investment in procurement, validation, and training. Forensic laboratories must validate the tools they use for the specific algorithms employed, a process that takes time and resources. Furthermore, historical case data will inevitably contain MD5 and SHA-1 hashes, requiring labs to maintain the capability to verify them. Interoperability with partner agencies or legal systems that have not yet fully migrated also plays a role.

In essence, a documented gap exists between the recommendations of the cryptographic community and standards bodies urging complete migration away from MD5 and SHA-1, and the operational practices within parts of the digital forensics community. This gap reflects differing risk assessments based on specific forensic use cases (integrity verification vs. signatures) and the significant practical challenges associated with universal, immediate migration across all tools, procedures, and legacy data. While the trend is undeniably towards SHA-2 and SHA-3, the handling of MD5 and SHA-1 remains a nuanced issue requiring careful consideration and clear documentation by forensic practitioners.

IV. Table: Overview of Hash Functions in Forensic Contexts

The following table summarizes the status, recommendations, and reliability considerations for major hash algorithms encountered in digital forensics, based on information from key standards bodies and forensic organizations discussed in this report. It aims to consolidate the guidance and provide clarity on the appropriate use of each algorithm family within the forensic context.

Hash Algorithm	Defining Standard / Guideline	Key Authority & Date	NIST Status / Timeline	Forensic Usage Recommendation & Reliability Notes	Key Sources (Snippet IDs)
MD5	RFC 1321	IETF 1992	Not FIPS Approved; Considered insecure due to collision attacks	**SWGDE (2019) / NFI (2018):** Acceptable *only* for integrity	

Hash Algorithm	Defining Standard / Guideline	Key Authority & Date	NIST Status / Timeline	Forensic Usage Recommendation & Reliability Notes	Key Sources (Snippet IDs)
			since 2004.	verification & file identification. **NOT** for signatures or security applications. Known collision vulnerability. Second preimage resistance considered less compromised historically, but use is discouraged. Fast.	
SHA-1	FIPS 180-4 (originally FIPS 180-1)	NIST 2015 (orig. 1995)	Deprecated 2011; Disallowed for signatures 2013; Practical collision 2017; **Phase-out by Dec 31, 2030.**	**NIST:** Not recommended, migrate ASAP. **SWGDE (2019) / NFI (2018):** Acceptable *only* for integrity verification & file identification. **NOT** for signatures. Known collision vulnerability. NFI switched preventively to SHA-256 in 2010 due to attack research.	
SHA-2 Family	FIPS 180-4	NIST 2015	**Approved.** Recommended for all applications	**NIST/INTERPOL/NFI:** Recommended standard for all	

Hash Algorithm	Defining Standard / Guideline	Key Authority & Date	NIST Status / Timeline	Forensic Usage Recommendation & Reliability Notes	Key Sources (Snippet IDs)
			requiring secure hashes.	forensic purposes (integrity, identification). **SHA-256 minimum** encouraged for interoperability. Considered secure.	
(SHA-224, **SHA-256**, SHA-384, SHA-512, SHA-512/224, SHA-512/256)					
SHA-3 Family	FIPS 202	NIST 2015	**Approved.** Secure alternative to SHA-2. No mandated transition *from* SHA-2 *to* SHA-3.	**NIST:** Approved for all applications. **NFI (2018):** Not yet common in forensics, SHA-256 considered sufficient. Secure alternative if needed/supported by tools.	
(SHA3-224, **SHA3-256**, SHA3-384, SHA3-512)					
(SHAKE128, SHAKE256)	*FIPS 202 (XOFs)*	*NIST 2015*	*Approved as XOFs; can be specialized to hash functions.*	*Less common in standard forensic integrity checks currently. May see use in specific cryptographic applications.*	

Note: The table reflects the status and recommendations based primarily on the provided

research snippets, focusing on NIST, SWGDE, NFI, and INTERPOL guidance relevant to digital forensics. "Acceptable" for MD5/SHA-1 refers specifically to the positions taken by SWGDE (2019) and NFI (2018) for integrity/identification tasks only, acknowledging these positions are debated and contrast with NIST's stronger deprecation stance. The trend is strongly towards SHA-256 and newer algorithms.

V. Conclusion and Forward-Looking Recommendations

A. Summary of Current State

The landscape of hash function usage in digital forensics is characterized by a significant, ongoing transition. Driven by the discovery of critical cryptographic vulnerabilities, particularly collision attacks, there is a clear and accelerating trend away from the legacy algorithms MD5 and SHA-1. Standards bodies, led by NIST, have established firm policies and timelines mandating the phase-out of SHA-1 and strongly recommending the use of the SHA-2 and SHA-3 families for all security-related applications, including those within the U.S. federal government. SHA-256 has emerged as the most widely adopted and recommended minimum standard globally, endorsed or exemplified by NIST, NFI, and INTERPOL for forensic work. However, the transition is not uniform across the global forensic community. Guidance from practitioner-focused groups like SWGDE and NFI has historically acknowledged the specific forensic use case of integrity verification, arguing that the less compromised second preimage resistance of MD5 and SHA-1 made them acceptable *for that specific task*, even while acknowledging their unsuitability for digital signatures. This nuanced position, while technically grounded in the differing requirements of forensic integrity checks versus broader cryptographic security, reflects the practical challenges of legacy systems, tool validation, and existing case data that forensic laboratories must manage. International bodies like ENFSI and INTERPOL emphasize the fundamental importance of using hashing for data integrity within a quality framework, increasingly pointing towards modern standards like SHA-256 as best practice.

B. Recommendations for Forensic Laboratories

Navigating this transitional period requires forensic laboratories to adopt proactive strategies that balance cryptographic best practices with operational needs and the requirement for forensically sound, admissible evidence. The following recommendations are proposed:

1. **Algorithm Selection:** Laboratories should adopt **SHA-256 as the minimum standard** for generating *new* hash values for all forensic purposes, including evidence integrity verification and file identification. This aligns with prevailing international recommendations and ensures robust protection against known vulnerabilities. Where tools and workflows support them, the stronger SHA-384, SHA-512, or SHA-3 family members can also be considered as secure alternatives. Generating new MD5 or SHA-1 hashes should be strictly avoided unless dictated by unavoidable legacy system constraints, and if done, the rationale must be thoroughly documented.
2. **Tool Validation:** The validation of forensic hardware and software is critical. Laboratories must rigorously test and validate their tools for the specific hash algorithms they intend to use, ensuring accuracy and reliability. Validation should confirm that the tool correctly

implements the chosen algorithm(s) (e.g., SHA-256) and produces accurate hash values consistent with reference implementations. All validation procedures and results must be documented.

3. **Handling Legacy Data:** Procedures must be in place to handle digital evidence that was originally hashed using MD5 or SHA-1. Laboratories must retain the capability to *verify* these legacy hash values. However, examiners must be acutely aware of the cryptographic limitations of these algorithms. Case notes, examination reports, and testimony should clearly and accurately state the specific hash algorithm used for verification and acknowledge its known limitations, particularly if the evidence is contested. Relying solely on an MD5 or SHA-1 hash match as proof of integrity may face increasing scrutiny.
4. **Policy and Training:** Clear, written policies should be established within the laboratory defining the approved hash algorithms for current work and the procedures for handling legacy hashes. Personnel should receive regular training covering the cryptographic properties of different hash functions, their known vulnerabilities, the distinction between collision and second preimage resistance in the forensic context, the laboratory's specific policies, and the critical importance of accurate documentation and reporting.
5. **Stay Informed:** The fields of cryptography and digital forensics are constantly evolving. Laboratories should designate personnel to monitor updates from standards bodies (e.g., NIST), forensic working groups (e.g., SWGDE, ENFSI), tool vendors, and the cryptographic research community regarding hash function security and best practices.

C. Future Outlook

The use of MD5 and SHA-1 in digital forensics will inevitably continue to decline, particularly as the NIST 2030 deadline for SHA-1 deprecation approaches, forcing tool vendors and laboratories reliant on FIPS validation to migrate. SHA-256 is likely to remain the workhorse algorithm for the foreseeable future, given its strong security track record and widespread implementation. The adoption of the SHA-3 family may increase gradually as tool support becomes more common and potentially accelerates if significant, practical weaknesses are ever discovered in the SHA-2 family, although NIST currently sees no need for such a mandatory switch.

Regardless of the specific algorithms used, the fundamental principles of digital forensic practice remain constant: meticulous documentation, rigorous validation of tools and methods, maintaining the chain of custody, and ensuring the integrity of digital evidence through scientifically sound procedures. While stronger hash algorithms provide greater assurance, they are only one component of a comprehensive quality management system necessary to produce reliable and admissible digital evidence. The future requires laboratories to manage the complexities of the ongoing cryptographic transition while upholding these core forensic principles.

Sources

1. Hash Functions | CSRC - NIST Computer Security Resource Center, https://csrc.nist.gov/projects/hash-functions

2. Cryptographic Hash Algorithm Competition | NIST, https://www.nist.gov/programs-projects/cryptographic-hash-algorithm-competition

3. nvlpubs.nist.gov,
https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-107r1.pdf

4. Recommendation for Applications Using Approved Hash Algorithms | NIST,
https://www.nist.gov/publications/recommendation-applications-using-approved-hash-algorithms

5. NIST Retires SHA-1 Cryptographic Algorithm,
https://www.nist.gov/news-events/news/2022/12/nist-retires-sha-1-cryptographic-algorithm

6. MD5 Archives - Schneier on Security -, https://www.schneier.com/tag/md5/

7. www.forensicinstitute.nl,
https://www.forensicinstitute.nl/binaries/forensicinstitute/documenten/publications/2018/02/13/forensic-use-of-hash-values-and-associated-hash-algorithms/Supplement-hashes-v2018_01a_English.pdf

8. Cybercrime Module 4 Key Issues: Standards and best practices for digital forensics,
https://sherloc.unodc.org/cld/en/education/tertiary/cybercrime/module-4/key-issues/standards-and-best-practices-for-digital-forensics.html

9. Supplement-hashes-v2018_01a_English | PDF | Computer File - Scribd,
https://www.scribd.com/document/848531537/Supplement-hashes-v2018-01a-English

10. MD5 and SHA1: Essential Hash Values in Digital Forensics - SalvationDATA,
https://www.salvationdata.com/knowledge/hash-value/

11. www.swgde.org,
https://www.swgde.org/wp-content/uploads/2024/04/2024-03-22-SWGDE-Best-Practices-for-Digital-Forensic-Video-Analysis-18-V-001-1.1.pdf

12. www.interpol.int,
https://www.interpol.int/content/download/13501/file/INTERPOL_DFL_GlobalGuidelinesDigitalForensics

13. Unique File Identification in the National Software Reference Library,
https://www.nist.gov/document/draft-unique-file-identification-nsrl-steve-mead-20060517

14. MD5 and SHA-1 Still Used in 2018 - Schneier on Security -,
https://www.schneier.com/blog/archives/2018/12/md5_and_sha-1_s.html

15. Digital Signature Standard (DSS) - NIST Technical Series Publications,
https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.186-4.pdf

16. Stateless Hash-Based Digital Signature Standard - NIST Technical Series Publications,
https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.205.pdf

17. Hash Functions | CSRC,
https://csrc.nist.gov/projects/hash-functions/nist-policy-on-hash-functions

18. Research Results on SHA-1 Collisions | CSRC - NIST Computer Security Resource Center,
https://csrc.nist.gov/news/2017/research-results-on-sha-1-collisions

19. Standards and Best Practices in Digital and Multimedia Forensics - Wiley Monthly Title Update and Image Download Site,
https://catalogimages.wiley.com/images/db/pdf/9781118640500.excerpt.pdf

20. The Future of Digital Evidence Authentication at the International Criminal Court,
https://jpia.princeton.edu/news/future-digital-evidence-authentication-international-criminal-court

21. Digital Investigation Techniques: A NIST Scientific Foundation Review,
https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8354.pdf

22. enfsi.eu,
https://enfsi.eu/wp-content/uploads/2016/09/1._forensic_examination_of_digital_technology_0.pdf

23. Recommendation for Key Management: Part 1 - General - NIST Technical Series Publications, https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-57pt1r5.pdf

24. Best Practice Manual - ENFSI,
https://enfsi.eu/wp-content/uploads/2022/12/FSA-BPM-002_BPM-for-Digital-Audio-Authenticity-Analysis.pdf

25. Best Practice Manual for Digital Image Authentication - ENFSI,
https://enfsi.eu/wp-content/uploads/2021/10/BPM_Image-Authentication_ENFSI-BPM-DI-003-1.pdf

26. GUIDELINES FOR DIGITAL FORENSICS FIRST RESPONDERS | Interpol,
https://www.interpol.int/content/download/16451/file/Guideline%20for%20First%20Responders%20Leaflet%20to%20be%20published%20on%20public%20INTERPOL.pdf

27. guide-first_responders-digital_devices_in_battlefield.pdf - the United Nations,
https://www.un.org/counterterrorism/sites/www.un.org.counterterrorism/files/guide-first_responders-digital_devices_in_battlefield.pdf

28. Projects - DFRWS, https://dfrws.org/projects/

29. An introduction to Hansken, https://www.hansken.nl/an-introduction-to-hansken

30. Digital forensics as a service: Stepping up the game - ResearchGate,
https://www.researchgate.net/publication/343821133_Digital_forensics_as_a_service_Stepping_

up_the_game

31. BKA - Homepage, https://www.bka.de/EN/Home/home_node.html

32. Technologies - BKA, https://www.bka.de/EN/OurTasks/SupportOfInvestigationAndPrevention/Technologies/technologies_node.html

33. Cybercrime - BKA, https://www.bka.de/EN/OurTasks/AreasOfCrime/Cybercrime/cybercrime_node.html

34. Forensic Science - BKA, https://www.bka.de/EN/OurTasks/Remit/CentralAgency/ForensicScience/forensicScience_node.html

35. 2nd Symposium on ICAO-Standard MRTDs, Biometrics and Security, https://www.icao.int/Meetings/AMC/MA/2006/MRTD_06/kinnegingseidel.pdf

36. Criminal Justice Information Services (CJIS) Security Policy - FBI.gov, https://le.fbi.gov/cjis-division/cjis-security-policy-resource-center/cjis_security_policy_v5-9-5_20240709.pdf

37. Guidance on digital forensics and protective monitoring specifications for producers of network devices and appliances, https://www.ic3.gov/CSA/2025/250204.pdf

38. Forensics — LE - FBI.gov, https://le.fbi.gov/science-and-lab/forensics

39. Computer Forensics - Department of Justice, https://www.justice.gov/sites/default/files/usao/legacy/2008/02/04/usab5601.pdf

40. Digital Evidence | FBI - LEB, https://leb.fbi.gov/articles/featured-articles/digital-evidence

41. Common Challenges in Cybercrime 2024 - Europol, https://www.europol.europa.eu/cms/sites/default/files/documents/Common_Challenges_in_Cybercrime_2024.pdf

42. Full article: Europol and cybercrime: Europol's sharing decryption platform - Taylor & Francis Online, https://www.tandfonline.com/doi/full/10.1080/14782804.2021.1995707

43. First report of the observatory function on encryption - Europol, https://www.europol.europa.eu/cms/sites/default/files/documents/final_report_of_the_observatory_function.pdf

44. Electronic evidence - a basic guide for First Responders - UNODC Synthetic Drug Strategy, https://syntheticdrugs.unodc.org/uploads/syntheticdrugs/res/library/cybercrime_html/Electronic_Evidence_A_Basic_Guide_for_First_Responders_ENISA.pdf

45. FIRST REPORT ON ENCRYPTION - Europol, https://www.europol.europa.eu/cms/sites/default/files/documents/EU_Innovation_Hub_First%20Report%20on%20Encryption.pdf

46. Advancing Digital Forensics investigations with the utilization of Cryptographic Inventories, https://cyberspaceproject.eu/advancing-digital-forensics-investigations-with-the-utilization-of-cryptographic-inventories/

47. AI and policing - The benefits and challenges of artificial intelligence for law enforcement - Europol, https://www.europol.europa.eu/cms/sites/default/files/documents/AI-and-policing.pdf

48. Forensic Hash Value Guidelines: Why Md5 and SHA1 should no longer be used and a recommendation for their replacement - ResearchGate, https://www.researchgate.net/publication/358979047_Forensic_Hash_Value_Guidelines_Why_Md5_and_SHA1_should_no_longer_be_used_and_a_recommendation_for_their_replacement

49. What would happen if SHA256 or MD5 broke down in one day ? : r/cybersecurity - Reddit, https://www.reddit.com/r/cybersecurity/comments/1bqnh1h/what_would_happen_if_sha256_or_md5_broke_down_in/

50. The Curious Case of MD5 - Hacker News, https://news.ycombinator.com/item?id=38861034

51. "The Impact of SHA-1 File Hash Collisions On Digital Forensic Imaging: " by Gary C. Kessler, https://commons.erau.edu/jdfsl/vol11/iss4/10/

52. SHA-1 Archives - Schneier on Security -, https://www.schneier.com/tag/sha-1/

53. Position on the Use of MD5 and SHA1 Hash Algorithms in Digital and Multimedia Forensics, https://www.swgde.org/position-on-the-use-of-md5-and-sha1-hash-algorithms-in-digital-and-multimedia-forensics/

54. Positions and Considerations - SWGDE, https://www.swgde.org/documents/positions-and-considerations/

55. SWGDE Minimum Requirements for Testing Tools used in Digital and Multimedia Forensics - National Institute of Standards and Technology, https://www.nist.gov/document/swgde-18-q-001-10-minimum-requirements-testing-tools-used-digital-and-multimedia-forensics

56. SWGDE 21-F-001-1.1, Best Practices for Acquiring Online Content, https://www.aafs.org/sites/default/files/media/documents/DigitalEvidence_SWGDE%2021-F-001-1.1_Checklist%20V1.xlsx 57. Digital Evidence Preservation: Considerations for Evidence Handlers - NIST Technical Series Publications, https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8387.pdf